The 11th Asian-Australasian Conference on Precision Agriculture (ACPA 11)
October 14-16, 2025, Chiayi, Taiwan

A Review on Structural Enhancements and Domain-Specific Adaptation of YOLO for Crop-Weed Recognition

Sohyeon Kang[1], Jaesung Park[1,2]*, Yeonsoo Kim[1,2], Jisu Song[1], Seokyung Park[1]

[1] Department of Bio-Industrial Machinery Engineering, Pusan National University, Miryang, Korea
[2] Life and Industry Convergence Research Institute, Pusan National University, Miryang, Korea

*Corresponding Author: jaesung.park@pusan.ac.kr

ABSTRACT

This review systematically summarizes YOLO-based weed detection models, focusing on two key directions: attention mechanisms that improve discrimination between visually similar vegetation and lightweight techniques that ensure real-time performance on limited hardware. A comparative analysis of improved YOLO variants highlights how structural optimizations improve detection, offering insights into efficient model design.

Keywords: Object Detection, YOLO, Weed Detection, Precision Agriculture.

INTRODUCTION

Deep learning has advanced precision agriculture, with weed management remaining critical as weeds compete with crops for resources. Conventional methods such as herbicides and mechanical weeding face environmental and crop damage risks, while early approaches like manual observation or rule-based image processing lacked scalability. CNNs improved robustness, and object detection models such as Faster R-CNN and especially YOLO further accelerated progress in weed detection. YOLO’s balance of accuracy and speed made it the most widely applied framework, yet weed detection remains difficult under complex conditions like variable illumination, occlusion, and crop–weed similarity. Recent studies emphasize attention mechanisms to improve feature discrimination under these challenging conditions, while lightweight design has emerged as a parallel direction to ensure real-time deployment in resource-constrained environments.

To our knowledge, no review has systematically analyzed YOLO-based structural enhancements specifically tailored for weed detection. This work addresses that gap by consolidating recent advancements, focusing on attention and lightweight strategies, and comparing YOLO-based models to provide insights for developing precision weeding systems.

MATERIALS AND METHODS

This review conducted a structured survey of YOLO-based weed detection models using Web of Science, Scopus, and Google Scholar. Searches combined the keywords “weed detection,” “weed identification,” “YOLO,” and “object detection,” limited to titles, abstracts, and author keywords in SCI/SCIE-indexed English papers from 2016 to 2025. About 250 papers were retrieved, with duplicates removed and around 50 highly relevant or frequently cited studies selected, supplemented by additional works where detailed context was required.

RESULTS & DISCUSSION

Structural improvement strategies for YOLO-based weed detection can be grouped into three categories: attention mechanisms, lightweight techniques, and other optimization approaches. Attention mechanisms reweight extracted features to enhance detection under crop–weed similarity, complex backgrounds, and variable illumination. Modules such as SE, CBAM, ECA, and Coordinate Attention, along with hybrid and 3D designs, have improved feature discrimination. However, improper integration may increase computational costs or degrade accuracy, making careful selection critical. Recent studies combine attention with multi-scale feature fusion, propose domain-specific lightweight modules, or apply Transformer-inspired self-attention to YOLO architectures. Lightweight techniques are equally important for real-time performance on limited hardware. Strategies include efficient backbones (ShuffleNet, MobileNet, HGNetV2), optimized necks (BiFPN, Lite-PAN), lightweight heads (LiteDetect, LDH), and efficient operators (DWConv, GhostConv, RepConv). Channel reduction, parameter sharing, and model compression techniques such as pruning, quantization, and knowledge distillation further reduce model size and FLOPs while preserving accuracy.

Approximately 20 YOLO-based weed detection models have been identified, most of which were built on YOLOv5 or YOLOv8. These models achieved improvements through lightweight design, attention-based modules, high-resolution feature layers, and advanced fusion strategies. YOLOv5 is typically chosen for efficiency, whereas YOLOv8 is preferred for higher accuracy, with the selection depending on research goals. Most models have been designed heuristically by combining proven modules, which yields practical gains but limits deeper optimization. As an alternative, Neural Architecture Search (NAS) has been proposed to automatically design task-specific architectures and uncover non-intuitive combinations. Yet NAS faces challenges, including implementation complexity and the risk of overfitting.

CONCLUSIONS

YOLO-based weed detection models continue to evolve through the integration of attention mechanisms and lightweight strategies. By enhancing feature discrimination and enabling real-time operation on limited hardware, these techniques have significantly improved performance under complex agricultural conditions. This review organizes the improvement strategies applied to various YOLO-based weed detection models and demonstrates how these techniques have been heuristically combined to enhance performance. It systematically presents the structural improvement trends of YOLO models specialized for weed detection.

ACKNOWLEDGEMENTS

This work was supported by Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry(IPET) through Open Field Smart Agriculture Technology Short-term Advancement Program Program(or Project), funded by Ministry of Agriculture, Food and Rural Affairs(MAFRA)(RS-2025-02313136)